

Zfp-37 Is a Member of the KRAB Zinc Finger Gene Family and Is

View metadata, citation and similar papers at core.ac.uk

brought

provided by Erasmus Univer

N. MAZARAKIS,* D. MICHALOVICH, A. KARIS, F. GROSVELD,¹ AND N. GALJART

MGC Department of Cell Biology and Genetics, Erasmus University, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands; and

*Laboratory of Neurobiology, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, United Kingdom

Received September 5, 1995; accepted January 18, 1996

The murine *Zfp-37* gene encodes a protein with 12 zinc fingers at its C-terminus (Nelki *et al.*, 1990, *Nucleic Acids Res.* 18: 3655; Burke and Wolgemuth, 1992, *Nucleic Acids Res.* 20: 2827–2834). Contrary to the published data, our Northern blot analysis demonstrates not only that the *Zfp-37* gene is expressed as 2.3, 2.6, and 4.2 kb mRNAs in testis, but also that there is a 3.7-kb message in the adult mouse brain. Using a partial cDNA as a probe, we have isolated a brain-specific *Zfp-37* cDNA clone of 3.3 kb, whose sequence was extended to full length using 5' end RACE. This revealed that the 3.7-kb mRNA is in fact a collection of transcripts with heterogeneous 5' ends. Comparison of cDNA and genomic sequences shows that the *Zfp-37* gene is spread over a region of ~20 kb and consists of six exons, the large 3' end exon containing the complete zinc finger domain, and 3' UTR. Our data show that the *Zfp-37* gene utilizes different promoters, alternative splicing, and differential polyadenylation to generate the distinct transcripts of brain and testis. Several protein isoforms are encoded by these mRNAs, some of which contain a truncated form of a conserved domain (Krüppel-associated box) found in other zinc finger genes. *In situ* hybridization analysis of postnatal brain sections indicates that the *Zfp-37* gene is expressed in all neurons of the central nervous system. Together, these results suggest that ZFP-37 is a transcriptional regulator predominantly present in postmitotic cells from two different lineages. © 1996 Academic Press, Inc.

INTRODUCTION

The zinc finger domain is one of a small number of nucleic acid binding motifs described to date (for a review see Harrison, 1991). It consists of 28–30 amino acids and contains four Cys and/or His residues that

coordinate the binding of a Zn²⁺ atom. In multi-zinc-finger proteins of the Cys₂His₂ subclass, the individual fingers are often connected by a highly conserved linker of seven amino acids, which was first described in the *Drosophila* transcription factor and segmentation protein *Krüppel* (Schuh *et al.*, 1986). A model was developed for these proteins in which each finger binds three nucleotides, with the C-terminus positioned on the 5' side of the DNA recognition sequence and the N-terminus facing the 3' end (Nardelli *et al.*, 1991; Pavletich and Pabo, 1991). Crystal structure determination of a number of zinc finger proteins bound to their cognate DNA sequences has shown this model to be incomplete, i.e., other modes of binding exist (Fairall *et al.*, 1993; Pavletich and Pabo, 1993). Moreover, not every finger in a multi-finger protein will actually contribute to DNA binding (e.g., Keller and Maniatis, 1992), and the ZF motif can be used for RNA rather than DNA binding (Koster *et al.*, 1991; Theunissen *et al.*, 1992).

Mammalian genomes contain several hundred ZF genes (Bellefroid *et al.*, 1989, 1993; Chowdhury *et al.*, 1987) that can be further categorized on the basis of other protein motifs that they have in common. One of these is called the *Krüppel*-associated box (KRAB) (Bellefroid *et al.*, 1991) or finger preceding box (FPB) (Rosati *et al.*, 1991). This conserved domain is present in about a third of all Cys₂His₂-ZF proteins (Bellefroid *et al.*, 1993) and may function as a transcriptional repression module (Margolin *et al.*, 1994; Pengue *et al.*, 1994; Witzgall *et al.*, 1994).

We are interested in identifying transcriptional regulators involved in neuronal differentiation pathways. We have focused our attention on the murine *Zfp-37* gene, for which partial cDNA sequences derived from testis cDNA libraries have been published (Burke and Wolgemuth, 1992; Nelki *et al.*, 1990). Both groups have shown that the *Zfp-37* gene contains 12 zinc fingers, located at the C-terminus of the protein and connected by the seven amino acid linker region, which is also found in *Krüppel*. Northern and Western blot analyses, as well as *in situ* hybridization and immunohistochemical data, suggested an exclusive role for the *Zfp-37* gene

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under Accession Nos. X89264–X89266.

¹ To whom correspondence should be addressed. Telephone: +31-10-4087169. Fax: +31-10-4360225.

in male germ cell development (Burke and Wolgemuth, 1992; Hosseini *et al.*, 1994). In contrast, we have found a brain-specific *Zfp-37* transcript in adult mouse brain. We show here that this 3.7-kb mRNA is heterogeneous at its 5' end and that it encodes various protein isoforms. Interestingly, some of these contain a truncated form of the KRAB domain. We also isolated a 35-kb genomic clone containing the complete *Zfp-37* gene and determined its genomic organization. *In situ* hybridization on postnatal mouse brain sections confirms that *Zfp-37* is expressed in the brain. More importantly, in the brain this putative transcriptional regulator is present in neuronal cells only, indicating that it might play a role in the development or maintenance of this cell type.

MATERIALS AND METHODS

Mice. Mice used to prepare RNA for Northern blots, RACE, and S1 analyses were of C57/BL6 or CBA origin. Mice used for making the brain cDNA library and *in situ* hybridizations were of the CBA/CA or CBA/C57BL strains.

RNA isolation, Northern blot, and S1 analyses. RNA was prepared according to the method of Auffray and Rougeon (1980). Total RNA (15 μ g) was fractionated on 0.8% agarose gels containing 0.66 M formaldehyde (Fourney *et al.*, 1988). RNA was blotted and hybridized according to standard procedures (Sambrook *et al.*, 1989). A 0.4-kb *EcoRI* fragment, labeled by random priming (Feinberg and Vogelstein, 1983), was used as a probe. RNA quality and loading were checked by staining the blot with methylene blue (Herrin and Schmidt, 1988). For S1 analysis, a genomic subclone, digested with *BspEI*, dephosphorylated, and subsequently digested with *BamHI*, was isolated from a low melting agarose gel and labeled at the *BspEI* site with T4 polynucleotide kinase (Sambrook *et al.*, 1989). S1 experiments were performed with 10–20 ng of probe and 50 μ g of total RNA from brain, testis, kidney, or tRNA as published previously (Antoniou *et al.*, 1988).

Isolation of cDNA and genomic clones. The mouse brain cDNA library was constructed in Lambda ZAPII (Stratagene), as suggested by the supplier, using RNA from mice sacrificed at day of birth (P0). Approximately 1×10^6 unamplified phage were immobilized onto nitrocellulose filters, which were baked and hybridized as described above. A gel-purified *EcoRI/XbaI* fragment from the short testis cDNA (Nelki *et al.*, 1990) was used as a probe. The filters were initially washed at low stringency, exposed wet to film, and then further washed and again exposed. Six strong and about 120 weaker duplicate double positives were detected with this procedure. Three strongly hybridizing clones (B2, B4, B6) were obtained after plaque purification rounds and were further characterized.

For the mouse genomic cosmid library, high-molecular-weight DNA (>200 kb) was prepared from undifferentiated AB-1 embryonic stem cells (McMahon and Bradley, 1990). DNA was partially digested with *MboI* and fractionated. Fragments larger than 30 kb were ligated into *BamHI*-digested pTCF cosmid arms (Grosveld *et al.*, 1982). The ligation mix was packaged using Gigapack Gold, essentially according to the protocol of the supplier (Stratagene). A 0.4-kb *EcoRI* and a 0.7-kb *EcoRI/XbaI* fragment, from the 5' and 3' ends, respectively, of the B4 cDNA clone described above, were used to screen replica nitrocellulose filters (about 1×10^6 colonies). One clone (PKO 3.3.4) hybridized to both probes. It was colony purified and checked for rearrangements against mouse genomic DNA using *Zfp-37*-specific probes. Nucleotide sequences of cDNA and genomic clones were determined using the dideoxy chain termination method (Sanger *et al.*, 1977) on double-stranded DNA (Murphy and Kavanagh, 1988). Conditions were modified for use with [α -³⁵S]dATP α S and Sequenase, as recommended by the supplier (USB). Sequence data were

analyzed with the GCG (Devereux *et al.*, 1984) and MacVector (Kodak IBI) software packages.

5' end RACE analysis. A modification of the original procedure for rapid amplification of cDNA ends (Frohman *et al.*, 1988) was followed to obtain full-length *Zfp-37* cDNAs. An oligonucleotide of 17 bases (complementary to nucleotides 347–364 in Fig. 2A) was hybridized to 100 μ g of total brain RNA from E14.5-day mouse embryos, and reverse transcription was carried out for 2 h at 41°C using super RT (HT Biotechnology Ltd.). Afterward, samples were passed through a centricon-100 filter (Amicon), washed, and dried on a speedvac. One-sixth of the material was used in a tailing reaction with dATP and calf thymus terminal transferase, according to the protocol of the supplier (Boehringer Mannheim). The reaction mix was then diluted 25 \times with water. Four percent of this sample was utilized in a first PCR (Saiki *et al.*, 1988) with a *Zfp-37*-specific 3' end primer (18 bases, complementary to nucleotides 332–349 in Fig. 2A) and two 5' end primers, i.e., 5'-CTATCTAGAGAGCTCGCG-GCCG-3' (a 22-mer adaptor) and the same sequence linked to 18 T's (40-mer). After a first cycle at 95°C (5 min), 75°C (4 min, addition of *Taq* polymerase), 45°C (2 min), and 72°C (15 min), 24 more cycles were performed at 94°C (45 s), 50°C (45 s), and 72°C (2 min) each. The PCR buffer composition was as suggested by the various suppliers. First round PCR material was diluted with water, and 0.1% was taken for a second round of PCR using a third *Zfp-37*-specific 3' end primer (35-mer, complementary to nucleotides 156–190 in Fig. 2A) and the 5' end adaptor described above. Thirty cycles at 94°C (45 s), 65°C (45 s), and 72°C (1 min) each were performed. After the PCR, samples were extracted with phenol:chloroform (1:1 v/v), passed over a centricon-100 filter, digested with proteinase K, extracted again, and ethanol precipitated. Samples were digested with *XbaI* and *PvuII*, gel purified, and subcloned. Nineteen subclones were sequenced and shown to be *Zfp-37* derived.

In situ hybridizations. Mouse tissues were dissected and frozen immediately on dry ice. Sections (10 μ m) were cut, transferred onto slides, air-dried and fixed in 4% paraformaldehyde. Slides were either used immediately or frozen at –20°C for small periods of time prior to hybridization. *In situ* hybridizations were performed at 64°C as previously described (Nieto *et al.*, 1992). An 800-bp *EcoRI* subclone containing the 5' end of the cDNA (Fig. 2A) and 300 bp of genomic upstream sequences were made into digoxigenin-labeled sense and antisense riboprobes, using T3 and T7 RNA polymerases and digoxigenin-11-UTP, as specified by the manufacturer (Boehringer Mannheim). Anti-digoxigenin antibodies, conjugated with alkaline phosphatase, were from the same supplier. Sections were mounted with Univert (BDH).

RESULTS

Brain-Specific *Zfp-37* Expression and cDNA Isolation

Pilot Northern blot experiments indicated that *Zfp-37* is expressed in the adult mouse brain as well as in testis (not shown). We therefore screened a mouse brain cDNA library with a *Zfp-37*-specific probe and picked up three clones with different insert sizes. These were subcloned, and the longest (B4) was sequenced completely. A 0.4-kb *EcoRI* fragment from its 5' end was next used to probe Northern blots with total RNA from various mouse embryonic and adult tissues. Only testis and brain express *Zfp-37* at easily detectable levels (Fig. 1). Two highly abundant mRNAs of 2.6 and 2.3 kb and a minor transcript of 4.2 kb are present in testis, as reported previously (Burke and Wolgemuth, 1992; Hosseini *et al.*, 1994). It is novel that a 3.7-kb message is easily detected in male and female mouse

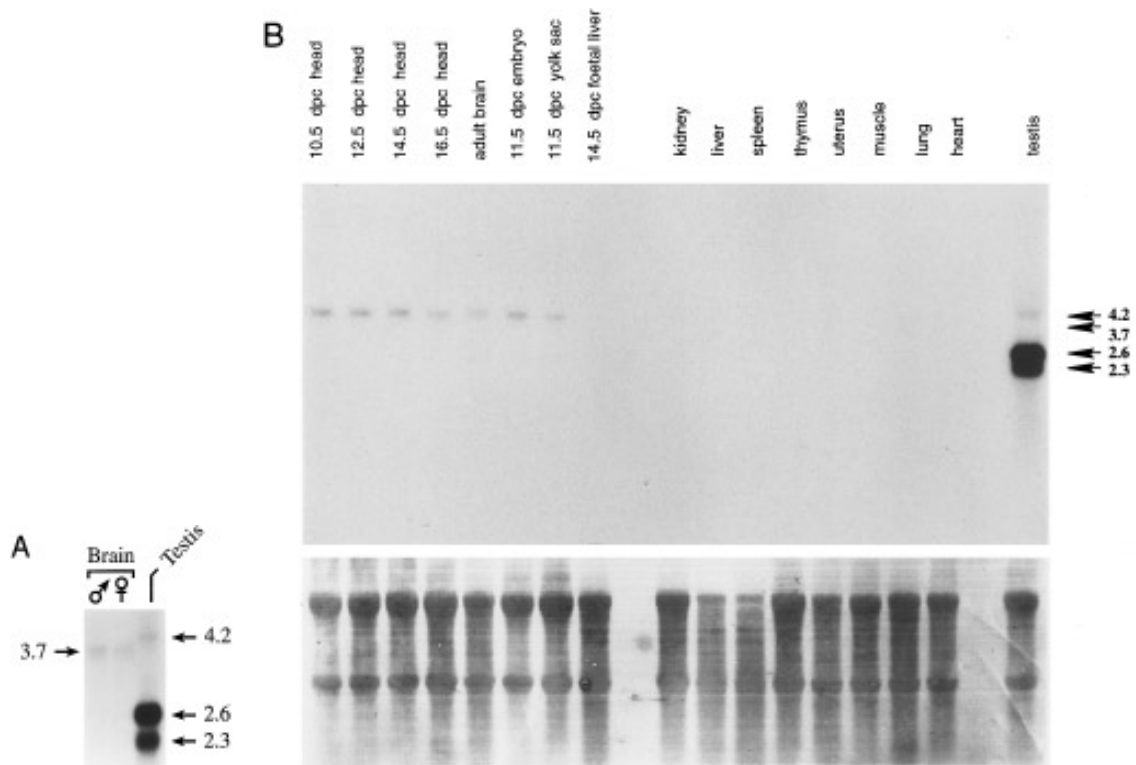


FIG. 1. Expression of *Zfp-37* in embryonic and adult mouse tissues. (A) A Northern blot with 15 μ g per lane of total RNA from adult brain (male and female) and testis was probed with a 0.4-kb *Eco*RI probe, derived from the 5' end of *Zfp-37* cDNA clone B4. After hybridization, the blot was washed and exposed to film for 1 week. Differences in length between the various *Zfp-37* transcripts are well visualized on this blot. Sizes are indicated and were calculated using RNA markers (0.3–7.4 kb). (B) (Top) A Northern blot with 15 μ g per lane of total RNA from various mouse tissues was hybridized as above. Exposure time was 3 days. The positions of the *Zfp-37* transcripts are indicated. (Bottom) The same blot was stained with methylene blue to measure quantity and quality of RNA. Note the slight degradation in liver and spleen samples.

adult brains (Fig. 1) and that low levels of this transcript are visible in heart, lung, uterus, thymus, liver, and kidney on long exposures of the autoradiograph (not shown). The 3.7-kb message is readily detectable in the developing head region from Day 10.5 of gestation onward (Fig. 1B). We have not tested whether the *Zfp-37* gene is expressed earlier in development. Probes derived from the zinc finger region gave similar results, as presented in Fig. 1, although additional weak signals around 1.4 and/or 2.4 kb were occasionally observed (not shown). Thus, our data show more than the previously published results (Burke and Wolgemuth, 1992; Hosseini *et al.*, 1994), probably because of the use of more sensitive probes in our case. Therefore, the most abundant expression of *Zfp-37* in the adult mouse is in testis, with a more moderate expression in the brain. Very low levels of transcript are detected in a number of other tissues.

The size of the longest cDNA insert (B4) was 3.3 kb, which appeared too short to account for a 3.7-kb mRNA. Therefore, a 5' end RACE analysis was performed to obtain a full-length cDNA. This yielded three different 5' ends. Most of the RACE clones (13/19) are of the so-called type Ia (Fig. 2A). Type Ia clones differ internally in that regions of nine and/or three nucleotides are

present in some clones, whereas they are absent in others. As shown below, this is due to alternative splicing. Note that these insertions do not cause a frameshift in the putative open reading frame (ORF). The sequences of the other two groups of 5' end clones, called Ib and II, are shown separately in Figs. 2B, and 2C, respectively. Thus, the 3.7-kb *Zfp-37* mRNA is heterogeneous at its 5' end. RT-PCR with novel sets of primers specific for exon Ia or Ib verified that both exons are present in *Zfp-37* mRNAs from adult brain and testis (not shown).

Comparison of the published *Zfp-37* sequences with the brain-specific cDNAs reveals several interesting facts. First, the 5' ends of brain- and testis-derived cDNAs differ completely in sequence from nucleotide position 131 in Fig. 2A (asterisk) going upstream. When a type Ia sequence is linked to the B4 cDNA, the ORF proposed by others (Burke and Wolgemuth, 1992) is extended toward the 5' end. Instead of ATG_{209–211}, it is now ATG_{86–88} that would be first encountered by the scanning ribosome. Assuming translation starts at this first ATG, protein isoforms of approximately 67 kDa are produced. These differ at their N-termini by one, three, or four amino acid residues, due to the variability in type Ia sequences that we have described

(AT) A TGC GAC TCT GAC TCC ATC GAC GAC CTG ACA AAG CAC GAG GCG CTG GAT CGG AGG AGT ACG GAC AAA GGC AGG GGT CGC CCG AGG GAG GAG ATG 1
S A S S D T S I E G D L T A K P E A G A L D R R S S A D K R R P P E G E (M) 1

100
GCT ATA TCC GAG CCT ACG GAA AGC GAT GCG GTG AGG GCA AAG GAA TGG GAG CAG CTG GAA CCT GTG CAG AGA GAT GTG TAC AAG GAT ACG 31
A T S B P G G A A G C G A V R A K E W E O L E F V G Q R D V Y K D T

200
AAG CTA GAG AAC TGC AGC AAT GCA GCC TCC ATG GGA AAT CAC GAT CCC AAA CAA GAC ATA GTC TCC GTG TTA GAA GAA GAG CCA TCA 61
K L E N C S N P A E (M) G N Q A D P K Q O D I V E V L E E E P S

300
TCG GGA AAG GGG AAA AAA GCC AGC CCA AGT AGT CTG AAA AAA ATA GCA AGG CCC AAG ACA GCA GGA AAT AGT GCA AAA CTC CAA CAA GAT 91
S G A K G K K A S C P A S T A L K K K A T A A G P K T A A G A C T A G T

400
GAT GAG CAT AGG GAG GAA AAG CAG AAG TCC CAA AGC AAA CTT ACT AAG GAA GTG ACA CTC AGG AAG AAA AGT TCC AAC AGC AAG AAA AGC 121
D E H R E E K Q A K (S) Q S A K T L T K E V T L R K K S S N S K K S

500
450 AGT GAG TAT GGT TTG TTG GAG AAG AAA AAT CTC CTA AAA CAC ACT CCT TCC GAG AAA AAA CTG CTT AAG TCC AGT TCC CGT GGG AAG 151
S E E Y T L L E N K A S L H S K H C A C T C T P C G E K K K C T L A K T C C T C C G T G G G A G K

600
AAC TGT AAT CAG AAT TCA GAC TCT CTG AAA AAG AAA CCT GAC ACA GCT AAT GAA CAC AGG AAA TCA CTC AGC CAT TCT GCA TCT GAT GTG 181
N S T N Q N S D S L K K K P D T A N E H R K S L S H S A S D V

700
AAC AAA GAT GAA ATT CCA ACT AAG AAA AAG AAA TGC GAC AAG TTA CCC AAC AAT AAG TTG TCT GAT AAA GGT GAG CAA AAC CAA ACT AGC AAA 211
N K D E I P T R A K K C D K L P N N K L S D K G D K N Q C S K

800
AAA TGT GAG AAA GTA TGC CGT CAT AGT GCA TCC CAT ACC AAG GAA GAC AAA ATT CAG ACC TGG GAG AAA CGG AAA TCA CAC TGC CGT ACT 241
C K E A V T C R H S S H T K E A G A A C I Q T G G G A A R K A S H C R T

900
CCA TCT AAA CCT GAA AAA GCC CCA GGT TCT GGG AAA CCT TAT GAA TGT AAC CAC TGT GGG AAG GTC CTC AGC CAT AAA CAG GGA CTC CTT 271
P S K P E K A P G S G K P T Y E C N H C G K V L S H K Q G L L

1000
900 GAC CAT CAA AGA ACT CAC ACT GGG GAG AAA P AT Y GAA TGT AAT GAA TGT GGT ATA GCT TCT AGC CAG AAG TCC CAC CTT GTT CTA CAT 301
D H Q T T H T G E K A P Y E C N E C G I A F S C A G K S H L V L H

1100
CAG AGA ACT CAC ACT ACT GGG GAA AAA CCA TAC GAG TGT GAA CAG TGT GGC AAA GCA CAC GGA CAT AAA CAT GCC CTC ACT GAC CAT CTA AGA 331
Q R T H T G E K P Y E C E Q C G K A H G H K H A L T D H L R

1200
ATC CAT CTG GAA GAA ACC TAC AAA TGT AAT GAA TGT GGC AAA ACG TTT AGA CAC AGC TCA AAC CTT ATG CAA CAC CTA AGA TCT CAC 361
I H T G E K P Y E C N E C G K A F K Y K T F R H S S L N L M Q C H L R S S H

1300
ACG GGT GAG AAG CCG TAT GAA TGT AAG GAA TGT GGC AAA TCC TTT AGA TAT AAT TCA TCT CTT ACT GAA CAT GTG AGA ACA CAC ACA GGT 391
T G E K P Y E C K E C G K S F R Y N S S L T E H V R T H T G

1400
GAA ATA CTA TAC GAA TGT AAG GAA TGT GGC AAA GCT TTT GAG TAT GGC TCA SCC CTG ACT AAA CAT ATG CGG ATT CAT ACA GGG GAG AAA 421
E I P Y E C N E C G K A F K Y K T F R H S S L N L M Q C H L R S S H

1500
1350 CCA TTT GAA TGT AAT GAA TGT GGG AAA ACT TTT AGC AAA AAG TCA CAC CTA GTT ATA CAT CAA AGA ACT CAT ACA AAG GAG AAA CCT TAT 451
P F E C N E C G K A T F S K K S H L V I H Q R T H I T K E K P Y

1600
AAA TGT GAT GAG TGT GGG AAA GCC TTT GGA CAT AGC TCA TCT CTT ACC TAT AAT GAA ACT CAT ACA GGT GAC TGC CCC TTT GAA TGT 481
K C D E C G K A F G H S S L T Y H M R T H T G E G D C C P F E A

1700
AAT CAA TGT GGT AAG GCC TTT AAA CAG ATT GAA GGC CTT ACC CAA CAG CAG AGA GTT CAC ACA GGG GAG AAA CCC TAT GAG TGT GTT GAA 511
N Q C G K A F K Q I E G L T Q H Q R V H T G E K P Y E C V E

1800
TGT GGG AAA GCC TTT AGT CAG AAG TCA CAC CTC ATC GTA CAG CAG AGA ACT CAT ACA GGT GAG AAA CCC TTT GAA TGT TAT GAG TGT GAG 541
C G K A F S C Q K S H L I V H Q R T H T G E K P F E C Y E C G

1900
AAA GCC TTC AAT GCA AAA TCA CAA CTT GTT ATT CAT CAG AGA TCC CAC ACT GGA GAG AAA CCC TAT GAA TGT ATT GAA TGT GGT AAA GCC 571
K A F N A K S Q L V I H Q R S H T G E K P Y E C I E C G K A

2000
1800 TTT AAG CAA AAT GCT TCT ACC AAT CAT ATG AAA ATT CAC TCA GAA GAA CAA TCT GAG GAA GAA GAT *A TGTAGGAAACCTGACACTGACTG 594
(F) K Q N A S L T K H M K I H S E Q S E E D

2100
GTTTGTGATTTTAACTTAAGAGTTGTTCTCAATTTGATGATGTGAGATATCTTTTTTTGAAGAAATCATTCCTGGTGATACATGAGAGAAATTTGAATATGGAATCTTTTACATTAAGATGG 594

2200
TAAATAAAATTTTAACTTTGATCCCAATCTTAAACAGAAAAATCTGATTTTTCACAGCTTATTATGAAATTTGTTCCATATTAGATTTAAATAACGAATCTTTTAAAGAGTGTAATGAGGA 594

2300
TATCATATACGTAAAGTCTCTTTTAAGCAAGTACAGAGATCTGAATTCCTACCTCCATAGGATTTGATAAAAGCTAGATGTGGGGGACGTCTCTGCAATCTCACAATTTGGGAAAGTTGAG 594

2400
ACAGAGAGCTCTGAGAACTCTGACCAACAGCACTCAATTTATTTGCAAGCATTAGGTCAAGAAAAAGAACTGTTTAAAAATAGGGTAGAGGGTTCTCGAGAAACAGTACTCAAG 594

2500
GTTGACACATGTACACACAGACACACAGACACACACACAGACACACAGATACACACACAGACACACACACAGACACACAGTTGTGAAAAATGTTA 594

2600
ACGAATTTAGAGCAACAAACAGTAAATATGACATCTCATGAACCTTTCTAGTGTGTGCTTAAAGCTCCACAGTATGATTATTAATTTTATTTATTTCTTATTGTCATCTACCGCTTGCTGT 594

2700
AAAATCAAGTGAGGGTGTAGGGTATCTCTGTGATTTGTCTAGCTTTCTTTTGTATTTCTCTGTTAGCTTTCTGAATTTAGTAGAAAACTTTGGGGAGTTGTTTGTGTTTGTGTGAGACAA 594

2800
ATCTCACCATATGCTCTGGCTAGCTTGTGATCTGTTACAGGCTGACCTTGAACCTTTTATGTAGACAGGCTGACCTTGAACCTCAGAGATATGCTGCTCTGCTCTGCCA 594

2900
AGTGCTGATTTTAAAGACTTCAGAAAACTTTAAGGGGTTTGTCTAGATTTTGTGTAAGAGGTTCAATAAAAGTTGAAAAACAGCTGAAAAATGTTGATAGTGAAGTATCTGTTAT 594

3000
ATTGTTAAGTATAGAAAAATGAGTTGTTTAAACATCTATATAGTATGTTCTCATCTTGACAGAGTGTGATATGTTGTCCTGCTACTCTGTGTACATATATATGTTTATCCCATATAT 594

3100
TTTAAATGATGATAGACCTTTTTTAACTATAGAAAAATTAATGTGGCTTCCACTGAACAAAGAGTGCAGGTTGAGGTTCTGACATGTTTATCATGTCATGTCATCTTAATCAAAATAG 594

3200
TATTATATAAAAATACATTCATATGTCAAATTAACAAATTTAGACAAATGAGTCTATTTTACAGATGTTATTTGAAAAATGACATAAGATTTCTATGCTCTTTATGAACAAAG 594

3300
AATTAAAGGACATGGCATTTGAACCTTGATTTGAATTCACAAAGAAATTTCCATAGAGCAGATTTTGTGTAACATGAGTCAATCTGTTAAACAAATTTTATTTCTCTTATTGCTA 594

3400
TTAAAACTAAAGATGAATGAAAAAATTAATGTGGCTTCCACTGAACAAAGAGTGCAGGTTGAGGTTCTGACATGTTTATCATGTCATGTCATCTTAATCAAAATAG 594

TTTGAGAACTCTGGCGTTCGTTGTTGAATAGTCTTTTGGGAAGGCCGGTGTCTTAAGGA
50
C AAGGATTTGGGGTATATATGGCCATATGGAGTATTAAGTGCCCGTGATATGGGGGCAT
100
ATGGAGTTGGGATGCTGTAACTTTCCGMAATCAGTATTTTGACGATCTGTCTAGAA
150
TGGGAGCAG

TTCATAGATTATCTGTGGCTTTTACTTAGAAGGAATGGGAGCAG

above. Since we cannot formally exclude the possibility that translation initiation takes place upstream of the first ATG, for example at a non-AUG codon, we have indicated the putative translated residues upstream of ATG₈₆₋₈₈ in italics in Fig. 2A. If the *Zfp-37* mRNAs start with type Ib or type II sequences, in-frame stop codons upstream of ATG₂₀₉₋₂₁₁ are present. As described by Burke and Wolgemuth (1992), only a single protein of 553 amino acids (M_r 62.6 kDa) would then be produced.

The first published *Zfp-37* testis cDNA (Nelki *et al.*, 1990) contains an insertion in the sequence that leads to a frameshift and an incorrect N-terminal domain. This mistake was discussed previously (Burke and Wolgemuth, 1992). Several other base differences are found when the various *Zfp-37* coding regions are compared, some of which cause amino acid alterations (circled residues in Fig. 2A). The mutations at nt 386 and 983 could represent polymorphisms; the others (nt 1667, 1684, and 1799) are probably mistakes in the published testis cDNAs.

The long 3' UTR of the B4 cDNA has several noteworthy features. It contains seven AUUUA elements in addition to the two present in the testis cDNA 3' UTRs (underlined in Fig. 2A). These sequences could shorten the half-life of *Zfp-37* mRNAs, as seen in other transcripts (for review see Decker and Parker, 1994). The 3' UTR also contains two putative immediate response boxes at positions 2577–2582 and 3094–3099 (IRBs; dotted boxes in Fig. 2A). These are DNA regulatory elements, which mediate the response to PDGF in the immediate-early JE gene (Freter *et al.*, 1992). IRBs are found in the 3' ends of many immediate-early genes. Like *Zfp-37*, some of these genes contain an incomplete IRB (6/7 bases correct). Finally, the B4 cDNA 3' UTR contains a CA-dinucleotide repeat of about 100 bases.

All ZFP-37 isoforms have 12 zinc fingers of the Cys₂His₂ type located at their C-termini. The fingers are connected by the *Krüppel*-like amino acid linker. A short, acidic tail of nine amino acids follows the last finger. The N-terminal portion of ZFP-37 is rich in charged/polar amino acids and contains a degenerate zinc finger motif near the first real finger. Such "rudimentary"

are also found in other members of this class of proteins (Attar and Gilman, 1992; Bellefroid *et al.*, 1993). Further inspection of the sequence reveals that ZFP-37 is actually a member of the KRAB/FPB family of zinc finger proteins (Bellefroid *et al.*, 1991; Rosati *et al.*, 1991); i.e., the type Ia cDNAs encode an almost complete KRAB domain, whereas the other cDNAs code for isoforms with only the degenerated "B" region. Together, our data suggest that the *Zfp-37* gene encodes various protein isoforms of the KRAB multi-zinc-finger family.

Organization of the *Zfp-37* Gene

The complicated structure of the *Zfp-37* cDNAs prompted us to determine the genomic organization of this gene. The *Zfp-37* gene could be isolated on a single cosmid clone. It covers roughly 20 kb and is located toward the 3' end of the 35-kb cosmid insert. Sequence analysis of genomic subclones and comparison to *Zfp-37* cDNAs suggests that the gene contains six exons, called Ia, Ib, It, II, III, and IV (Fig. 3A). Exons Ia and Ib are colinear with the 5' RACE clone types Ia and Ib, respectively. Exon It (t stands for testis) was found downstream of exon Ib and comprises 17 bases. These nucleotides align perfectly to the extreme 5' end (without the first 7 bases) of one of the published *Zfp-37* testis cDNAs (Burke and Wolgemuth, 1992) and are followed by a putative splice donor in the genomic sequence (Fig. 3B). The first seven nucleotides of the testis cDNA (Burke and Wolgemuth, 1992) are Lambda-ZAP-derived.

The 5' end RACE analysis suggests that transcription of the *Zfp-37* gene starts from three different promoters and that some of the primary transcripts are alternatively spliced. We have summarized these putative events in Fig. 3C. The major transcription initiation site(s) (11/19 RACE clones start here) is 5' to exon Ia. The occurrence of type Ib (2/19) and type II (4/19) RACE clones suggests that transcription can initiate at two other promoters, one just downstream of exon Ia and the other 5' to exon II. Alternative splicing of exon Ia to exon II occurs at two donor sites, which are

FIG. 2. Brain-specific *Zfp-37* cDNA sequences. (A) Composite sequence of 5' end RACE-derived sequences linked to the 3.3-kb mouse brain cDNA clone (B4), isolated from a P0 cDNA library. The RACE analysis yielded three types of clones; the most common one (type Ia; 13/19 clones) is linked to B4. Most type Ia clones start 2 nucleotides apart (brackets at 5' end). Two shorter type Ia clones start at nt 107 and 108 (vertical arrows). The 5' ends of the B4 cDNA and the short testis clone are marked, as is the 3' end of the testis cDNAs (Burke and Wolgemuth, 1992; Nelki *et al.*, 1990). The point of 5' end divergence between long testis cDNA (Burke and Wolgemuth, 1992) and brain cDNAs is marked by an asterisk. The two potential polyadenylation sites (solid lines) and IRBs (dotted lines) are boxed. The nine AUUUA elements are underlined (wavy line). Amino acid numbering is indicated to the right. The first two methionine residues are circled. The amino acid sequence upstream of the first methionine is italic to indicate that it may be translated. Alternatively spliced amino acid residues (VRA/K) are underlined, as are the 12 zinc fingers at the C-terminus of ZFP-37. Amino acid substitutions due to differences in the various sequences are circled. The KRAB domain is shaded. (B) Two type Ib cDNA clones derived from the 5' end RACE analysis. The sequence of the longer clone is given; the sequence of the second clone is colinear with the first (dotted underlining) up to the vertical arrow (splice site). The shorter clone has an insertion of 3 nucleotides at this position. Downstream of the arrow, cDNA clones become colinear. In the long clone, the presence of three short ORFs is indicated. (C) Four type II clones derived from the 5' end RACE analysis. The sequence of the longest clone is given; the others are indicated with dotted lines. Vertical arrows mark splice acceptor sites of exon II (see below).

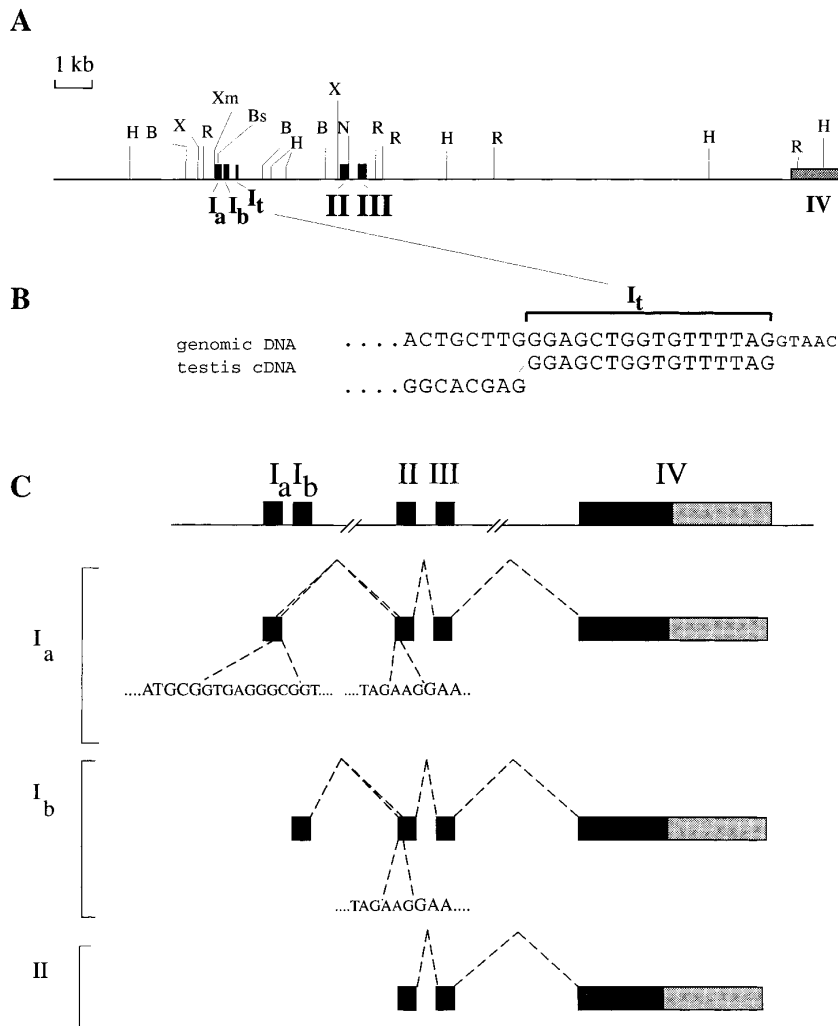


FIG. 3. Genomic organization and transcription pathways of the *Zfp-37* gene. (A) A cosmid with an insert of 35 kb was isolated from a mouse genomic library. The *Zfp-37* gene is located toward the 3' end of this insert, on an area of about 20 kb. A partial restriction enzyme map of this region is shown. The six exons are Ia, Ib, It, II, III, and IV; their positions are indicated by boxes (nucleotide positions of exon-intron boundaries within the *Zfp-37* cDNA can be found in the EMBL database). Exon IV contains the 3' UTR (shaded box). H, *Hind*III; B, *Bam*HI; X, *Xba*I; R, *Eco*RI; Xm, *Xma*I; Bs, *Bsp*EI; N, *Nco*I. (B) The 5' end of one of the two testis cDNAs (Burke and Wolgemuth, 1992) was found to align perfectly to a genomic sequence of 17 bases, which was therefore termed exon It (t stands for testis). The first seven nucleotides of the cDNA were not homologous to the genomic sequence and were later found to be lambda-ZAP-derived. At the 3' end of exon It there is a splice donor site, and the remainder of the testis cDNA aligns to exons II, III, and IV. This makes it likely that exon It is genuine, but its 5' border has yet to be defined. (C) Putative *Zfp-37* transcription initiation sites and modes of alternative splicing of *Zfp-37* transcripts in the brain. The 5' end RACE clones map to three different positions within the gene, suggesting that transcription initiation takes place in front of exons Ia, Ib, and II. Splicing of exon Ia to exon II occurs via two splice donor and two splice acceptor sites located 9 and 3 nucleotides apart, respectively (see sequences below the Ia-II splicing pathway). Splicing of Ib to II makes use of the same two acceptor sites.

nine nt apart, and at two splice acceptor sites, which are three nt apart. Splicing of exon Ib to exon II again makes use of the two acceptor sites at the 5' end of exon II (Fig. 3C).

Preliminary sequence analysis of a genomic 450-bp region 5' to and covering exon Ia shows it to be GC-rich (66%) and to contain 40 CpG dinucleotides (not shown), indicative of a CpG island (Bird, 1986). These usually colocalize with the 5' end of genes, suggesting that the composite cDNAs depicted in Fig. 2A are full length. This was confirmed by S1 nuclease protection

with a genomic fragment of 870 bp, labeled at a *Bsp*EI site (position 79, Fig. 2A) and covering the complete CpG island and upstream sequences. The results of this experiment show two sites of transcription initiation in the brain (arrows in Fig. 4), one of which colocalizes with the 5' end of the type Ia cDNA clones. In the testis, transcription initiation is spread over an area of more than 200 bases, with the bulk taking place some 250 bases upstream of the brain sites (bracketed in Fig. 4). S1 analysis with a probe closer to the testis transcription initiation sites gives similar results (not shown).

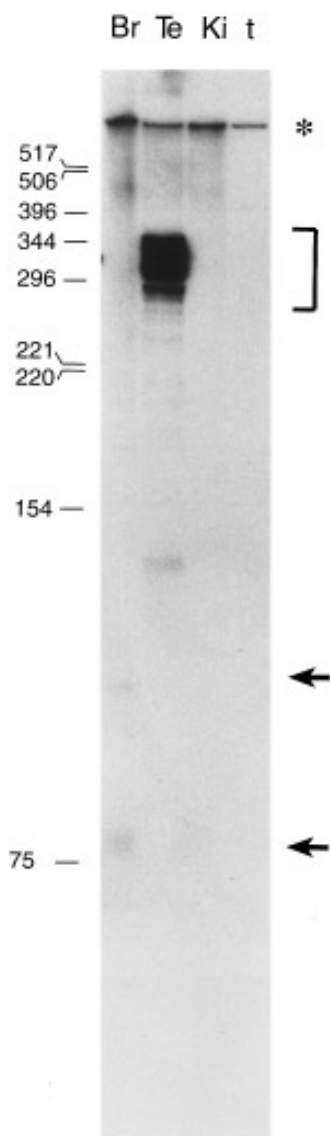


FIG. 4. *Zfp-37* transcription initiation at exon Ia as measured by S1 analysis. A genomic probe spanning the region from the *Bsp*EI site in exon Ia toward a more 5' *Bam*HI site (Fig. 3A) was used in a S1 nuclease protection experiment with 50 μ g of total RNA from brain (Br), testis (Te), kidney (Ki), or tRNA (t). S1-protected fragments were separated on a sequencing gel (marker DNA sizes are indicated to the left). The probe is visible in all lanes (asterisk). Multiple transcription initiation sites are present in testis (bracket), whereas there are only two weakly visible sites in brain (arrows).

Thus, transcription initiation from within the presumptive CpG island gives rise to ZFP-37 isoforms with the KRAB-A domain. Initiation outside the CpG island yields isoforms without this region.

At the 3' end of exon IV, the genomic sequence is colinear with the different cDNAs until the start of the respective poly(A) stretches, demonstrating the presence of two polyadenylation sites in the *Zfp-37* gene. Preliminary experiments with a probe between these two sites fail to show the 2.6- and 2.3-kb messages, whereas the 3.7-kb mRNA is detected (not shown).

Therefore, the first of the two polyadenylation sites is predominantly used in testis, whereas the second is predominantly used in brain.

Expression of *Zfp-37* in Mouse CNS

In situ hybridization was used to examine the spatial distribution of *Zfp-37* in the developing embryo at 12.5 d.p.c. and in postnatal mouse brains. In the E12.5 embryo, the strongest hybridization is observed in the germinal neuroepithelium, predominantly in the diencephalon, mesencephalon, and myelencephalon, although structures outside the CNS are also positive (Fig. 5A). The signal is present in the ventricular zone, indicating that the gene is expressed in multipotent precursor cell populations (Figs. 5C and 5D). In the neonatal brain (P3, P6, and P15), *Zfp-37* is expressed by a subset of neurons in the developing cerebellum (Fig. 5J), hippocampus, and neocortex (not shown). In the cerebellum, stronger hybridization is observed in Purkinje cells than in the granular layer at these early stages of development.

In adult brain, *Zfp-37* expression appears to be panneuronal (Figs. 5E–5H and 5L–5N), although the differential hybridization within the various neuronal populations is still observed. In the hippocampus, there is intense hybridization with pyramidal neurons and dentate granules (Figs. 5E–5G), but little or no hybridization in regions where glia are present, such as in the fimbria in the hippocampus, suggesting either that this gene is exclusively expressed in neurons or that it is also expressed in glia but at very low levels. In the olfactory bulb, *Zfp-37* expression is observed in the mitral and granule cell layers (Fig. 5N). Strong hybridization is also observed throughout the cortex (Fig. 5I). The hybridization signal is mainly cytoplasmic, but nucleoli are also stained (Fig. 5G). Parallel control experiments with the sense probe were negative (Figs. 5B and 5K), except for a weak nucleolar staining in every cell. These data suggest that *Zfp-37* is on early in precursors that give rise to the nervous system and that as the brain develops, *Zfp-37* expression becomes restricted to the neuronal but not the glial cell population. In neurons, *Zfp-37* expression is differentially regulated within distinct populations. *Zfp-37* expression is not exclusive to neurons, since the Northern blot data show that *Zfp-37* is also present outside the nervous system in both embryonic and adult organisms.

DISCUSSION

The *Zfp-37* gene was previously shown to encode a protein with 12 zinc fingers at its C-terminus. Two testis *Zfp-37* cDNAs have thus far been published, a short truncated one of 1.4 kb (Nelki *et al.*, 1990), which contains a sequencing mistake leading to an ORF with an incorrect N-terminal domain, and a longer clone of 1.9 kb (Burke and Wolgemuth, 1992), which was postu-

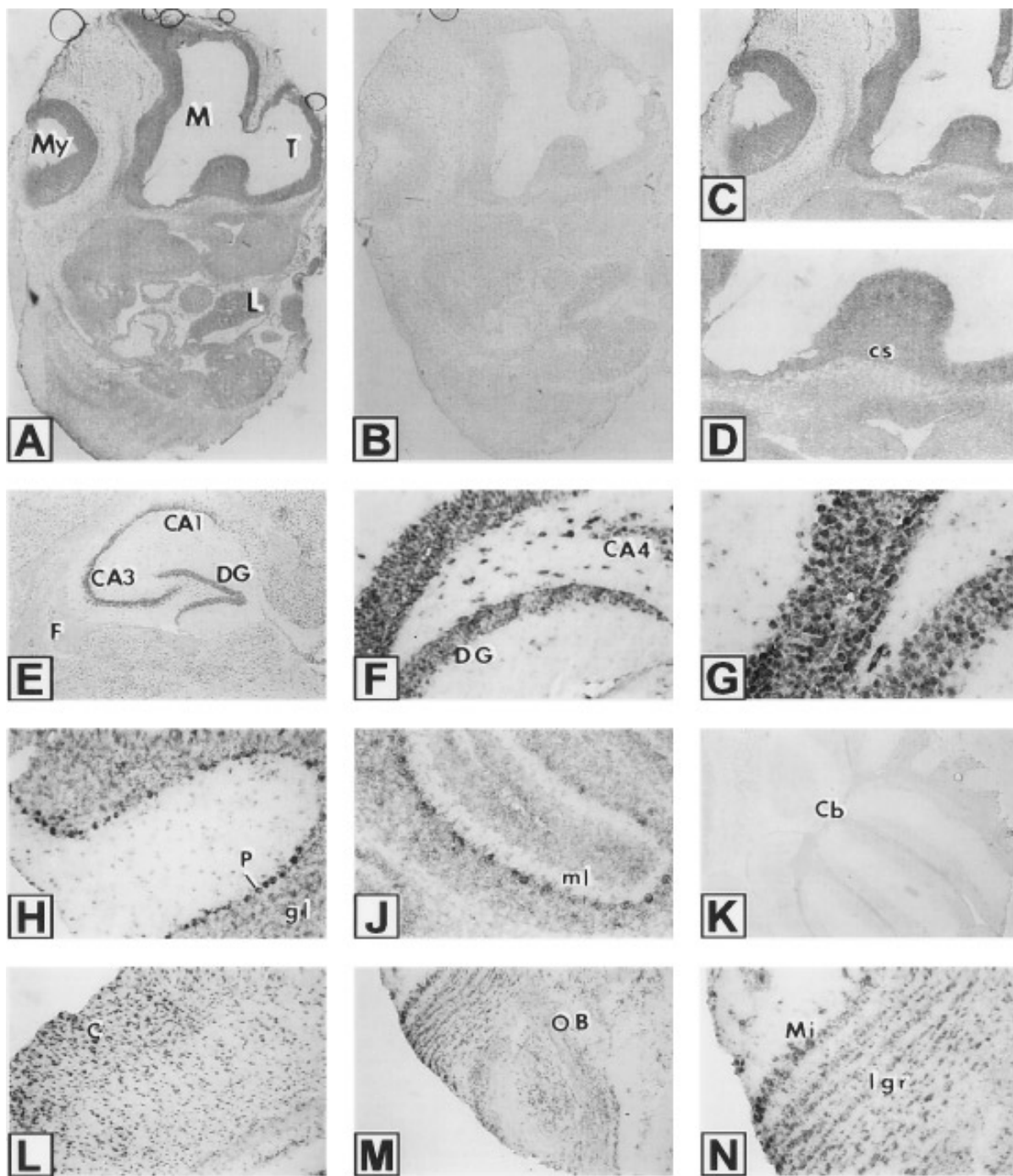


FIG. 5. *In situ* hybridization analysis of *Zfp-37* in mouse embryo and adult mouse brain. Frozen tissue sections of 10 μ m were fixed on slides and hybridized with digoxigenin-labeled riboprobes. Label was visualized using anti-digoxigenin antibodies conjugated to alkaline phosphatase. (A) E12.5-day embryo, antisense probe. The most intense signal is observed in ventricular layers of the head (germinal neuroepithelium), while there is also some hybridization above background in the rest of the trunk. (B) Same stage embryo, sense probe. Only a general background labeling is observed. (C) Magnified picture of E12.5-day embryo showing the hybridization of the antisense probe around the third and fourth ventricles. (D) Magnification of the presumptive striatum in the same embryo. (E) Adult hippocampus, antisense probe. Strong hybridization in the pyramidal and dentate granule neuronal layers. Little hybridization is present in the fimbria, suggesting neuronal specificity for the signal. (F) Labeled dentate granules (10 \times). (G) Labeled dentate granules (20 \times). (H and J) Adult and P6 cerebellum, antisense probe, showing strong hybridization of Purkinje neurons and weaker staining of granules. (K) Adult cerebellum, sense probe. (L) Cerebral cortex, antisense probe. (M) Olfactory bulb, antisense probe. (N) Magnification of olfactory bulb, showing hybridization in mitral cells and internal granules. T, telencephalic vesicle; M, mesencephalic vesicle; My, myelencephalic vesicle; L, liver; cs, corpus striatum; F, fimbria; CA1, CA3, CA4, pyramidal fields of hippocampus; DG, dentate granule layer of hippocampus; P, Purkinje cell layer; gl, granular layer of cerebellum; ml, molecular layer of cerebellum; Cb, cerebellum; C, cortex; OB, olfactory bulb; Mi, mitral cells; Igr, internal granular layer.

lated to be full length and to give rise to a protein with a M_r of 62 kDa. In the adult mouse, *Zfp-37* expression was detected only in testis, with the mRNA being localized to round and elongating spermatids. ZFP37 was therefore speculated to be a transcriptional regulator of spermiogenesis (Burke and Wolgemuth, 1992). We present here the cloning and characterization of brain-specific *Zfp-37* cDNAs as well as the isolation of a cosmid of 35 kb containing the complete *Zfp-37* gene. Three experimental approaches, i.e., Northern blot analysis, RT-PCR, and *in situ* hybridization, all confirm that the *Zfp-37* gene is expressed not only in testis but also in brain. We suggest that the composite cDNAs described in this paper are full length and represent the 3.7-kb *Zfp-37* transcript found on Northern blots. As far as the other transcripts that appear to be testis-specific are concerned, S1 analysis demonstrated that in the CpG island, testis transcription initiates roughly 250 bases upstream of brain transcription. mRNAs generated at one of these upstream sites and identical to the 3.7-kb message until the first polyadenylation site would have predicted lengths of about 2.6 kb, if a poly(A) tail of 300 bases is included in the calculation. This size correlates well with the 2.6-kb message found in testis. The 1.9-kb testis cDNA published previously (Burke and Wolgemuth, 1992) could represent the 2.3-kb mRNA found on Northern blots, but its transcription start needs to be mapped. Finally, the 4.2-kb mRNA probably represents another zinc finger gene (data not shown).

Those *Zfp-37* mRNAs containing exon Ia code for protein isoforms of about 67 kDa, which is 5 kDa more than was predicted previously (Burke and Wolgemuth, 1992). A KRAB domain is present in these ZFP-37 isoforms (Bellefroid *et al.*, 1991; Rosati *et al.*, 1991). The KRAB domain has been divided into a highly conserved A region and a degenerate B region. *In vitro* studies suggest that the KRAB domain acts as a DNA binding dependent repressor of transcription and that the A region is necessary and sufficient to obtain this effect (Margolin *et al.*, 1994; Pengue *et al.*, 1994; Witzgall *et al.*, 1994). Obviously, the functional significance of the ZFP-37 KRAB region needs to be established, but its presence suggests that ZFP-37 is a DNA rather than a RNA binding protein. In this context, it is noteworthy that fingers 2 and 10 of ZFP-37 are almost identical and could bind the same sequence. In contrast, fingers 1 and 3 have amino acid substitutions in a conserved phenylalanine (F to L in finger 1; F to H in finger 3). This could compromise their capacity to bind nucleic acid.

Expression of ZFP-37 isoforms is controlled at various levels. First, transcription of the gene appears to start in three regions, with the bulk of transcripts originating in the presumptive CpG island in which exon Ia is located. The neuronal nitric oxide synthase (nNOS) gene also has two closely linked promoters, like the Ia and Ib presumptive promoters of *Zfp-37*, and

one nNOS promoter is embedded in a CpG island (Xie *et al.*, 1995). The functional significance for *Zfp-37* is that mRNAs initiated at the CpG island encode ZFP-37 isoforms with a KRAB domain, whereas the other transcripts do not. Thus, a switch in promoter usage leads to the production of different proteins. This is not the case for nNOS.

A second level of control of expression could be mRNA stability, as suggested by the presence of nine AUUUA elements in the 3' UTR of the *Zfp-37* gene. Turnover of steady-state levels of 3.7-kb *Zfp-37* mRNAs may be much higher than that of 2.6- and 2.3-kb transcripts, which have only two AUUUA elements in their 3' UTRs. A very similar situation exists in the CREM gene, which was shown to be highly expressed in testis compared to brain due to nine AUUUA elements in the brain 3' UTR (Foulkes *et al.*, 1993). The presence of these elements and of two incomplete IRBs (Freter *et al.*, 1992), indicates that *Zfp-37* could be a member of the immediate-early gene family.

In addition to control exerted at the level of transcription and mRNA stability, regulation of mRNA translation is yet another means by which ZFP-37 protein levels may be kept in check within a cell. In the case of the type Ib *Zfp-37* mRNAs, ribosome translation at the correct ATG (209–211) can only take place if three short upstream ORFs are first bypassed. It has been shown for the yeast transcription factor *GCN4* that such downstream initiation only occurs when cells are starved (Hinnebusch, 1994). Therefore, it will be particularly interesting to follow the expression of type Ib *Zfp-37* mRNA; perhaps it can be linked to a specific cellular event or to a subtype of cells.

Together, the data suggest a model of ZFP-37 function in which the expression of the gene is carefully regulated at the level of transcription as well as translation. It is tempting to speculate that the different protein isoforms fulfill functions that have to do with the regulation of transcription of genes in postmitotic cells, since *Zfp-37* expression in the adult mouse is most abundant in round and elongating spermatids and in neurons. It remains to be determined whether ZFP-37 function is the same in both tissues and throughout development and whether the protein is a transcriptional activator and/or repressor. It could also be that *Zfp-37* gene products are involved in maintaining higher order chromatin structure, for example the opening or closing of gene loci. This would resemble the insulator function of *suppressor of hairy wing*, Su(Hw) (Roseman *et al.*, 1993), a protein with 12 zinc fingers found in *Drosophila melanogaster* (Parkhurst *et al.*, 1988).

ACKNOWLEDGMENTS

We thank Dr. Daniel Nelki for providing the 1.4-kb *Zfp-37* testis cDNA, Vangelis Stavropoulos for assisting with the subcloning and sequencing of the brain *Zfp-37* cDNA clones, Ton Verkerk for per-

forming the RT-PCR, Mirko Kuit for photographic work, and Drs. Angela Nieto and David Wilkinson for providing *in situ* hybridization protocols and advice. N.M. was supported in part by an ISRT fellowship, D.M. by a MRC studentship, and N.G. by long-term EMBO and Dutch Royal Academy of Sciences (KNAW) fellowships. This research was supported by the MRC (UK) and the Erasmus University Rotterdam.

REFERENCES

- Antoniou, M., deBoer, E., Habets, G., and Grosveld, F. (1988). The human beta-globin gene contains multiple regulatory regions: Identification of one promoter and two downstream enhancers. *EMBO J.* 7: 377–384.
- Attar, R. M., and Gilman, M. Z. (1992). Expression cloning of a novel zinc finger protein that binds to the c-fos serum response element. *Mol. Cell. Biol.* 12: 2432–2443.
- Auffray, C., and Rougeon, F. (1980). Purification of mouse immunoglobulin heavy-chain messenger RNAs from total myeloma tumor RNA. *Eur. J. Biochem.* 107: 303–314.
- Bellefroid, E. J., Lecocq, P. J., Benhida, A., Poncelet, D. A., Belayew, A., and Martial, J. A. (1989). The human genome contains hundreds of genes coding for finger proteins of the Kruppel type. *DNA* 8: 377–387.
- Bellefroid, E. J., Poncelet, D. A., Lecocq, P. J., Revelant, O., and Martial, J. A. (1991). The evolutionarily conserved Kruppel-associated box domain defines a subfamily of eukaryotic multifingered proteins. *Proc. Natl. Acad. Sci. USA* 88: 3608–3612.
- Bellefroid, E. J., Marine, J. C., Ried, T., Lecocq, P. J., Riviere, M., Amemiya, C., Poncelet, D. A., Coulie, P. G., de Jong, P., Szpirer, C., et al. (1993). Clustered organization of homologous KRAB zinc-finger genes with enhanced expression in human T lymphoid cells. *EMBO J.* 12: 1363–1374.
- Bird, A. P. (1986). CpG-rich islands and the function of DNA methylation. *Nature* 321: 209–213.
- Burke, P. S., and Wolgemuth, D. J. (1992). Zfp-37, a new murine zinc finger encoding gene, is expressed in a developmentally regulated pattern in the male germ line. *Nucleic Acids Res.* 20: 2827–2834.
- Chowdhury, K., Deutsch, U., and Gruss, P. (1987). A multigene family encoding several “finger” structures is present and differentially active in mammalian genomes. *Cell* 48: 771–778.
- Decker, C. J., and Parker, R. (1994). Mechanisms of mRNA degradation in eukaryotes. *Trends Biochem. Sci.* 19: 336–340.
- Devereux, J., Haeblerli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 12: 387–395.
- Fairall, L., Schwabe, J. W., Chapman, L., Finch, J. T., and Rhodes, D. (1993). The crystal structure of a two zinc-finger peptide reveals an extension to the rules for zinc-finger/DNA recognition. *Nature* 366: 483–487.
- Feinberg, A. P., and Vogelstein, B. (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 132: 6–13.
- Foulkes, N. S., Schlotter, F., Pevet, P., and Sassone-Corsi, P. (1993). Pituitary hormone FSH directs the CREM functional switch during spermatogenesis. *Nature* 362: 264–267.
- Fourney, R. M., Miyakoshi, J., Day, R. S., and Paterson, M. C. (1988). Northern blotting: Efficient RNA staining and transfer. *Focus* 10: 5–7.
- Freter, R. R., Irminger, J. C., Porter, J. A., Jones, S. D., and Stiles, C. D. (1992). A novel 7-nucleotide motif located in 3′ untranslated sequences of the immediate-early gene set mediates platelet-derived growth factor induction of the JE gene. *Mol. Cell. Biol.* 12: 5288–5300.
- Frohman, M. A., Dush, M. K., and Martin, G. R. (1988). Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA* 85: 8998–9002.
- Grosveld, F. G., Lund, T., Murray, E. J., Mellor, A. L., Dahl, H. H., and Flavell, R. A. (1982). The construction of cosmid libraries which can be used to transform eukaryotic cells. *Nucleic Acids Res.* 10: 6715–6732.
- Harrison, S. C. (1991). A structural taxonomy of DNA-binding domains. *Nature* 353: 715–719.
- Herrin, D. L., and Schmidt, G. W. (1988). Rapid, reversible staining of Northern blots prior to hybridization. *BioTechniques* 6: 196–197.
- Hinnebusch, A. G. (1994). Translational control of GCN4: An in vivo barometer of initiation-factor activity. *Trends Biochem. Sci.* 19: 409–414.
- Hosseini, R., Marsh, P., Pizzey, J., Leonard, L., Ruddy, S., Bains, S., and Dudley, K. (1994). Restricted expression of a zinc finger protein in male germ cells. *J. Mol. Endocrinol.* 13: 157–165.
- Keller, A. D., and Maniatis, T. (1992). Only two of the five zinc fingers of the eukaryotic transcriptional repressor PRDI-BF1 are required for sequence-specific DNA binding. *Mol. Cell. Biol.* 12: 1940–1949.
- Koster, M., Kuhn, U., Bouwmeester, T., Niefeld, W., el-Baradi, T., Knochel, W., and Pieler, T. (1991). Structure, expression and in vitro functional characterization of a novel RNA binding zinc finger protein from *Xenopus*. *EMBO J.* 10: 3087–3093.
- Margolin, J. F., Friedman, J. R., Meyer, W. K., Vissing, H., Thiesen, H. J., and Rauscher, F., III. (1994). Kruppel-associated boxes are potent transcriptional repression domains. *Proc. Natl. Acad. Sci. USA* 91: 4509–4513.
- McMahon, A. P., and Bradley, A. (1990). The Wnt-1 (int-1) proto-oncogene is required for development of a large region of the mouse brain. *Cell* 62: 1073–1085.
- Murphy, G., and Kavanagh, T. (1988). Speeding-up the sequencing of double-stranded DNA. *Nucleic Acids Res.* 16: 5198.
- Nardelli, J., Gibson, T. J., Vesque, C., and Charnay, P. (1991). Base sequence discrimination by zinc-finger DNA-binding domains. *Nature* 349: 175–178.
- Nelki, D., Dudley, K., Cunningham, P., and Akhavan, M. (1990). Cloning and sequencing of a zinc finger cDNA expressed in mouse testis. *Nucleic Acids Res.* 18: 3655.
- Nieto, M. A., Gilardi-Hebenstreit, P., Charnay, P., and Wilkinson, D. G. (1992). A receptor protein tyrosine kinase implicated in the segmental patterning of the hindbrain and mesoderm. *Development* 116: 1137–1150.
- Parkhurst, S. M., Harrison, D. A., Remington, M. P., Spana, C., Kelley, R. L., Coyne, R. S., and Corces, V. G. (1988). The *Drosophila* su(Hw) gene, which controls the phenotypic effect of the gypsy transposable element, encodes a putative DNA-binding protein. *Genes Dev.* 2: 1205–1215.
- Pavletich, N. P., and Pabo, C. O. (1991). Zinc finger-DNA recognition: Crystal structure of a Zif268–DNA complex at 2.1 Å. *Science* 252: 809–817.
- Pavletich, N. P., and Pabo, C. O. (1993). Crystal structure of a five-finger GLI-DNA complex: New perspectives on zinc fingers. *Science* 261: 1701–1707.
- Pengue, G., Calabro, V., Bartoli, P. C., Pagliuca, A., and Lania, L. (1994). Repression of transcriptional activity at a distance by the evolutionarily conserved KRAB domain present in a subfamily of zinc finger proteins. *Nucleic Acids Res.* 22: 2908–2914.
- Rosati, M., Marino, M., Franze, A., Tramontano, A., and Grimaldi, G. (1991). Members of the zinc finger protein gene family sharing a conserved N-terminal module. *Nucleic Acids Res.* 19: 5661–5667.
- Roseman, R. R., Pirrotta, V., and Geyer, P. K. (1993). The su(Hw) protein insulates expression of the *Drosophila melanogaster* white gene from chromosomal position-effects. *EMBO J.* 12: 435–442.
- Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R.,

- Horn, G. T., Mullis, K. B., and Erlich, H. A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239: 487–491.
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning: A Laboratory Manual", 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74: 5463–5467.
- Schuh, R., Aicher, W., Gaul, U., Cote, S., Preiss, A., Maier, D., Seifert, E., Nauber, U., Schroder, C., Kemler, R., *et al.* (1986). A conserved family of nuclear proteins containing structural elements of the finger protein encoded by Kruppel, a *Drosophila* segmentation gene. *Cell* 47: 1025–1032.
- Theunissen, O., Rudt, F., Guddat, U., Mentzel, H., and Pieler, T. (1992). RNA and DNA binding zinc fingers in *Xenopus* TFIIIA. *Cell* 71: 679–690.
- Witzgall, R., O'Leary, E., Leaf, A., Onaldi, D., and Bonventre, J. V. (1994). The Kruppel-associated box-A (KRAB-A) domain of zinc finger proteins mediates transcriptional repression. *Proc. Natl. Acad. Sci. USA* 91: 4514–4518.
- Xie, J., Roddy, P., Rife, T. K., Murad, F., and Young, A. P. (1995). Two closely linked but separable promoters for human neuronal nitric oxide synthase gene transcription. *Proc. Natl. Acad. Sci. USA* 92: 1242–1246.